AUTOMATED DETECTION OF CORONAL MASS EJECTIONS

DISSERTATION

SUBMITTED
TO
DEPARTMENT/SCHOOL PHYSICAL SCIENCES
DOON UNIVERSITY, DEHRADUN
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE OF

MASTERS
IN
PHYSICS

BY

DEEPAK KATHAIT
(17PH-04)

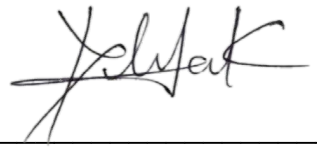
DEPARTMENT OF PHYSICS,
SCHOOL OF PHYSICAL SCIENCES,
DOON UNIVERSITY, DEHRADUN
UTTARAKHAND (INDIA)

2022

Declaration

I declare that the work presented in the Dissertation entitled 'Automated Detection of Coronal Mass Ejections' being submitted to the Department of Physics, School of Physical Sciences, Doon University, Dehradun for the award of Master in Physics is my original research work.

The Dissertation embodies the results of investigations, observations, and experiments carried out by me. I have neither plagiarized any part of the dissertation nor have submitted same work for the award of any other degree/diploma anywhere.



(Signature)

DEEPAK KATHAIT

(17PH-04)

Certificate

This is to certify that the Dissertation entitled "AUTOMATED DETECTION OF CORONAL MASS EJECTIONS' submitted by DEEPAK KATHAIT has been done under my supervision. It is also certified that the work in this Dissertation embodies original research and hard work of the candidate.

The assistance and support received during the course of investigation and all the sources of literature have been fully acknowledged.

Supervisor/Guide

Dr. Vaibhav Pant
Scientist-C
Aryabhatta Research Institute of Observational Sciences (ARIES)
Manora Peak, Beluwakhan
Nainitaal, Uttarakhand, India

Co-Supervisor/Guide

Dr. Vikas Sharma
Assistant Professor
Department of Physics
School of Physical Sciences
Doon University,Dehradun

Head of Department

Dr. Himani Sharma
Assistant Professor,
Department of Physics
School of Physical Sciences
Doon University, Dehradun

Acknowledgement

This work could not have been possible without the gracious opportunity provided by **Prof. Dipankar Banerjee,** Director, Aryabhatta Research Institute of Observational Sciences (ARIES) for which I am greatly indebted.

I express my profound gratitude to my project supervisor **Dr Vaibhav Pant**, Scientist at ARIES, for his exemplary guidance, monitoring, his constructive criticism and for providing a conductive research atmosphere throughout the course of this project Without his guidance, I would not be able to give concrete form to my ideas.

I am grateful to my co-supervisor **Dr Vikas Sharma,** Assistant professor, Doon University for he was the one who instilled the basics of coding within me in my bachelor years.

My gratitude extends to **Mr Ritesh Patel**–SRF, ARIES, Nainital, without his guidance I would have been unable to complete my project. His altruistic attitude has impacted me significantly. His willingness to help with the smallest of problems even at late hours shows his dedication to his craft. I am deeply grateful for his time, energy and genuine support during the dissertation.

I am also thankful to my **HOD, Dr Himani Sharma** for always believing in me, pushing me forward and for her unwavering support. She has been a constant source of inspiration for me.

I owe a debt of gratitude to my friend, **Mr Abhas Pradhan** for his invaluable efforts and support during the accomplishment of the thesis. This work also serves as a token for the great people I've met and wonderful time I've spend at ARIES. I'll always treasure these memories.

Finally, I offer my sincere regards to my family and friends for maintaining a faith in me.

Please accept my deepest gratitude.

Deepak Kathait

Abstract

Coronal mass ejections (CMEs) are sporadic expulsion of huge plasma structures, magnetic field embedded in those plasma structures and energy from the Sun. CMEs are an important part of coronal and interplanetary dynamics. They cause massive transient disturbances by injecting hefty amounts of plasma and magnetic flux into the heliosphere. CMEs are recognized to be the critical drivers of space weather and are a major source of solar energetic particles (SEPs). Thus it is necessary to keep track of such events.

With the advancement in modern technology, we now have an almost 24/7 observation of the sun. This generates massive amounts of data and with the upcoming launch of India's AditiyaL1 the data is only going to increase. Therefore there is a need for automated detection algorithms which will greatly reduce human effort and resources.

This work focuses on developing an algorithm for automated detection of CMEs using the Generalized Hough Transform. Although there are already automated detection catalogues such as CDAW and CACTUS, this work hopes to create a more efficient technique.

Table of Contents

Table of Figures

1. Introduction

1.1 The Sun

The crown jewel of our solar system, the progenitor of life on earth, the Sun is a rather average, main sequence star (yellow dwarf) at the core of our solar system. It's a giant ball of extremely hot ionized gas called plasma heated due to the thermonuclear reactions in its heart. It has a radius of 700,000 Km and mass 330,000 times that of earth [6]. Almost three-fourths of sun's mass is hydrogen, remainder is Helium with minor quantities of heavier elements. The Hertzsprung-Russell diagram (HR diagram), developed in early 1900s by Ejnar Hertzsprung and Henry Norris Russell are an important contraption to survey the stellar evolution. It constructs a plot between temperature and luminosity of the stars. If these two characteristics of a star are known they then can be used as coordinates in the HR diagrams (Fig. 1.1) which then can give an insight into the evolutionary phase of the star.

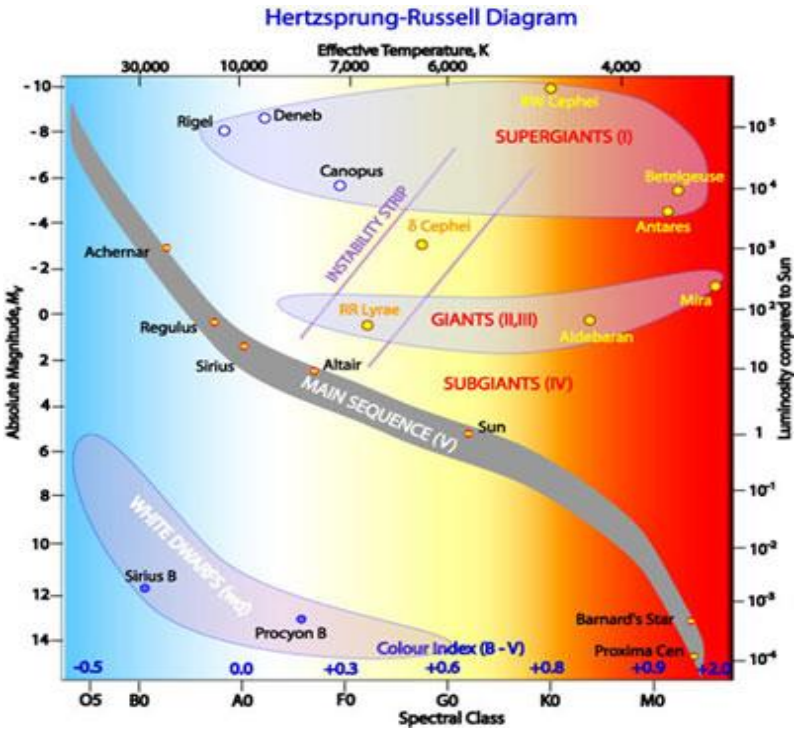


Fig. 1.1: The HR Diagram. The main sequence (grey), which runs from the top left (hot, brilliant stars) to the bottom right (cool, dim stars), is by far the most noticeable element. White dwarfs are situated below the main sequence, whereas the giant branch and supergiant stars are located above it.

***Credit**: R. Hollow, CSIRO.*

Our sun is in the main sequence, with a luminosity of 1 and a temperature of 5400K. The stars spend 90% of their lives in main sequence, burning hydrogen to produce helium.

1.2 The Structure of the Sun

The Sun is not a solid entity. It does not have a surface or discrete boundaries like other rocky celestial bodies. The sun comprises of many regions. In radially outwards direction, the core, radiative zone and convective zone constitute the solar interior. It is then followed by photosphere, chromsphere, transition region and corona which makeup the solar atmosphere. Outside corona is the continuous outflow of particles termed as solar wind. Let us look into each layer separately.

Core: The core is the furnace of the sun. It is the hottest part of the sun with temperature of almost 15 million degree C [8]. Hydrogen is used up to form Helium by the process of nuclear fusion. 700 million tons of hydrogen is converted to helium every second. It is also the densest part of the sun with density up to 150g/cm^3. Temperature and density both dwindles as we move away from the center of sun. The proton-proton chain, often known as the pp chain, is a three-step mechanism that our sun uses to fuse hydrogen to helium [5].

Step 1- Two protons strike together to give birth to a positron, deuteron and an electron neutrino.

$$p + p \rightarrow {}^{2}D_1 + e^{+} + v_e$$

The positron then annihilates with an electron to produce two gamma rays

$$p + p + e \rightarrow {}^{2}D_1 + e^{+} + v_e + 1.442\ MeV$$

Step 2- A proton bumps with the deuteron to form Helium 3 nucleus and gamma ray

$$^{2}D_{1} + {}^{1}H_{1} \rightarrow {}^{3}He_{2} + \gamma + 5.493Mev$$

Step 3- Two Helium-3 nucleus collides to form a Helium-4 nucleus and two protons by the process called the p-p I branch.

$$^{3}He_{2} + {}^{3}He_{2} \rightarrow {}^{4}He_{2} + 2\ {}^{1}H_{1} + 12.859MeV$$

In total 26.732MeV energy is released. There are three other possible ways of converting He-3 to He-4 called the p-p II, III, IV branches. The CNO also contributes to around 1% of the solar output.This reaction has fueled the sun for 4 billion years and will keep it up for another 5 billion.

The Radiative zone: It is the layer adjacent to the core and is characterized by transfer of energy generated in the core by the process of thermal radiation, hence the name radiative zone. The energy generated by the core is transported by photons that get absorbed and re-emitted by the ions present, many times over as they pass through the radiative zone. Even though photons travel with the speed of light, it takes them several hundred thousand years to cross the radiative zone. It stretches from 0.25 to 0.7 solar radii. It is the thickest layer of the sun. Radially outwards the density decreases from $20g/cm^3$ to $0.2g/cm^3$, while the temperature drops off from $7{,}000{,}000^{o}$ C to $2{,}000{,}000^{o}$ C.

Tachocline: It is the layer that lies between the convective and radiative zones. This is the location when the radiative zone's uniform rotation abruptly changes to the convective zone's differential rotation. Here, the layers slide one over the other. Currently, it is believed that a magnetic dynamo in this zone generates the magnetic field of the sun.

The Convective zone: At 0.7 solar radii the temperature has dropped to the point that transport of energy by radiation is not feasible instead the density of the plasma becomes low

enough to favor the convective currents. In this region, the temperature drops to around 2,000,000^{o} C, which is chilly enough for heavier ions like C, N, O, and Ca to retain part of their electrons. The material also becomes more opaque, making it more difficult for radiation to get through. Convective currents start to form as the medium boils due to the trapped heat. The hot substance expands, causing its density to decrease and it to climb above. The cycle continues as the cooled material drops below, absorbs heat, and rises once more. Granules and supergranules near the surface are evidence of the convective motion.

The Photosphere: This is the sun's visible surface (bright yellow); underneath it, it turns opaque to visible light.. It is about ten to hundreds of kilometers thick and has an average temperature of 6000K. This layer is responsible for emitting the white light which we can see with our naked eyes. The top of the thermal columns of the convective zone appear as granules in the photosphere.

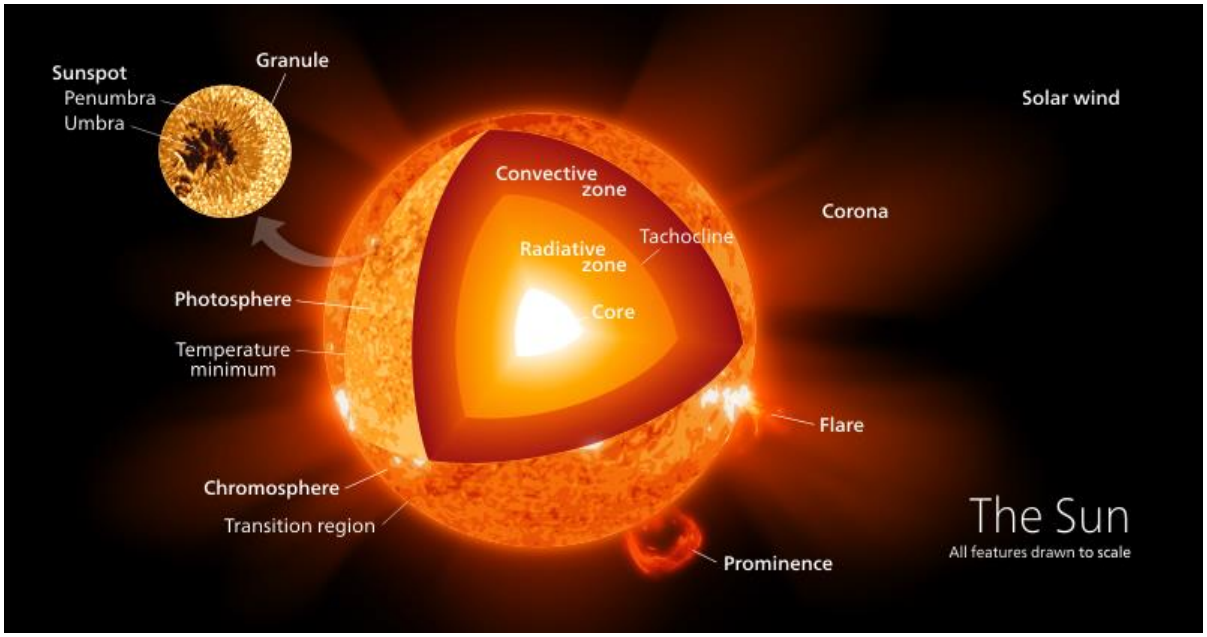


Figure 1.2: The above image shows different layers and features of the sun

Credits- https://en.wikipedia.org/wiki/File:Sun_poster.svg

Chromosphere: The temperature in this layer above the photosphere has grown from 6000^{o} C to 20,000^{o} C. It is 2500 kilometers thick on average [6].The word chromosphere means color-sphere as this layer has a reddish color due to H-α emission. This layer is hotter than the visible photosphere which is puzzling as the temperature should decrease as you move away from the core.

Transition Region: It is the region where the chromosphere becomes corona. The temperature abruptly rises from thousands of degrees to millions of degrees. Hydrogen gets ionized at these temperatures and instead of hydrogen, ions like C4+, O4+, and Si4+ dominate the light released by the transition region. The light emitted by these ions lie in the ultraviolet region that is only accessible from space. It is estimated to be 60 miles thick [8]

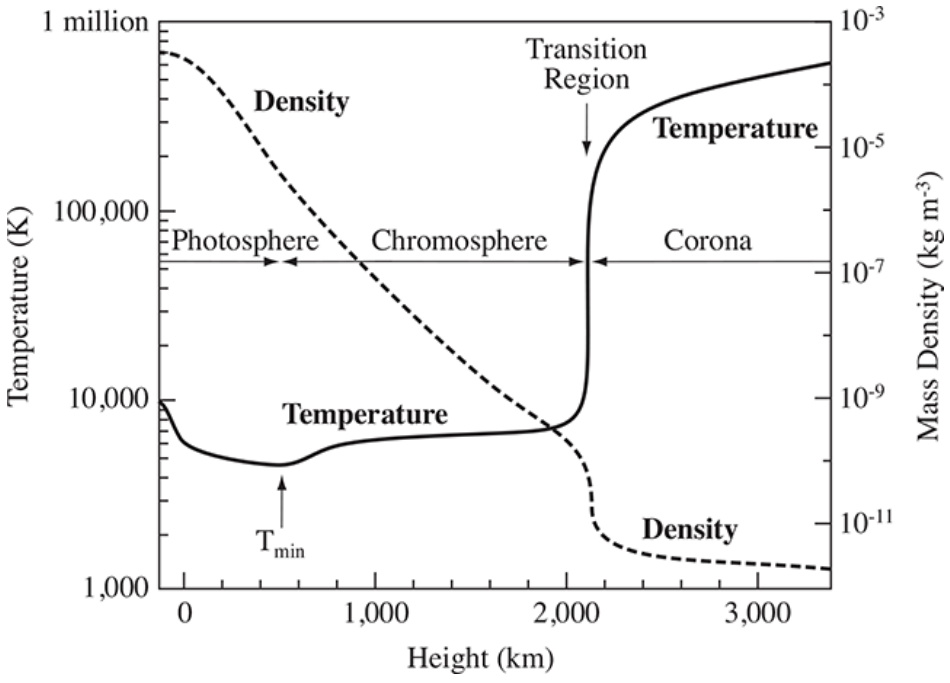


Figure 1.3: The figure plots profiles of mass density (dashed line) and temperature (solid line) in the solar atmosphere (courtesy of Eugene Avrett, Smithsonian Astrophysical Observatory.

Corona: The upper atmosphere of our sun is called the corona. It appears as white crown during a total solar eclipse. Early observations of the corona exposed bright emission lines which did not relate to any know material or transition. This led to the belief of presence of a new element called “coronium”. Later it was discovered that the transition line belonged to highly ionized states of iron (Fe X, Fe XIV) and calcium which requires very high temperature. The temperature here reaches few million Kelvin which like the chromosphere is again puzzling and famously called the Coronal heating problem.

1.3 Solar Cycle

Samuel Heinrich Schwabe, a German Astronomer for 17 years from 1826 to 1843 would scan the sun and its sunspots in order to find any planets inside Mercury’s orbit. Although he did not discover any new planets, it ultimately led to the discovery of periodicity of the sunspots with progression of the solar cycle. This later led to the discovery of 11 year cycle of the sun known as the solar cycle or the Schwabe cycle in his honor. Any solar cycle starts with solar minima, where the solar activity is low and as the time progresses the activity slowly starts to ramp up and we finally reach the solar maxima where the activity is at its highest. The sun’s atmosphere changes from calm and quiet to violently active. From the solar maxima the activity starts to decay and we once again reach the state of solar minima. After every 11 years, during a solar maxima the polarity of the sun’s pole also flips and after 22 years they go back to their original polarities. This is called the Magnetic Solar Cycle.

You must be wondering what I mean by the solar activity. Well to put it in simpler terms it is basically the commotion on the sun’s surface driven by the magnetic fields. Sunspots number and area, solar flares, CMEs, high speed solar winds etc are all a measure of the solar activity. As the

cycle progresses from minima to maxima, their rate of occurrences of these features also increases and vice versa. These activities are of vital importance as they determine the heliospheric conditions and are critical drivers of the space weather.

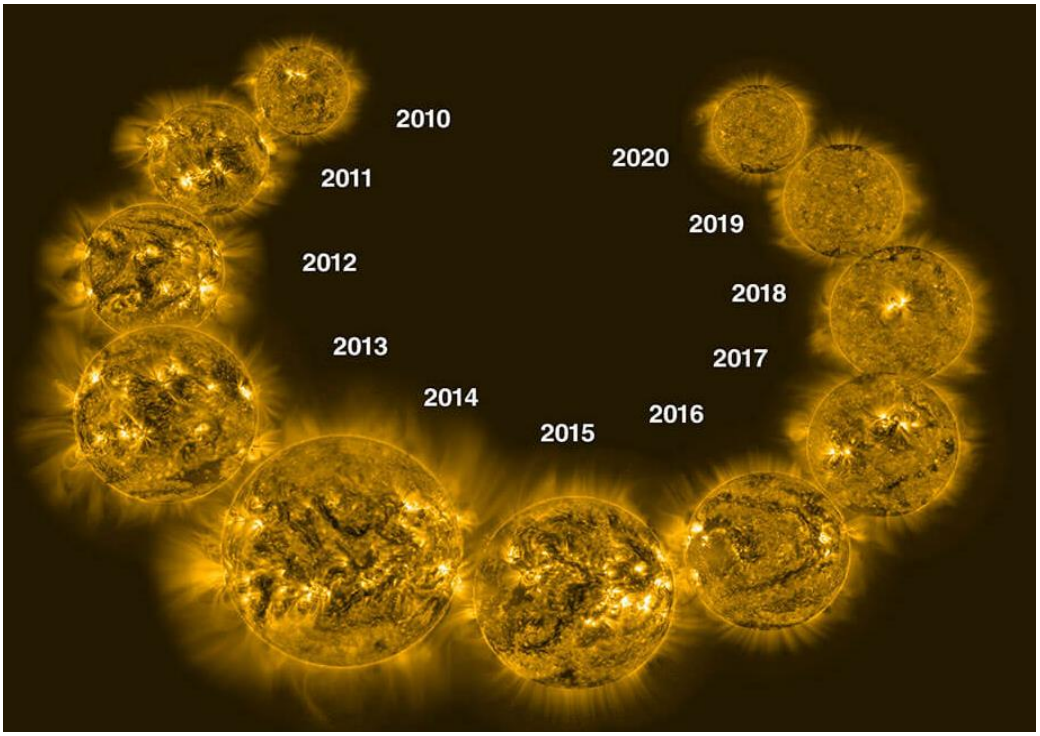


Fig. 1.4:Evolution of the Sun in extreme ultraviolet light from 2010 through 2020, as seen from the telescope aboard Europe's PROBA2 spacecraft. 2014-15 corresponds to solar maxima.
Credit: Dan Seaton/European Space Agency (Collage by NOAA/JPL-Caltech)

1.4 Coronal Mass Ejections (CME)

CMEs are sporadic expulsion of bulky plasma structures embedded with magnetic field lines from the sun into the heliosphere or simply space (Fig. 1.5). An average CME has mass of the order 10^{11}-10^{12}Kg [11] and can achieve speeds of several thousand kilometers per second. When a CME enters the interplanetary space, typically at a distance greater than 50 solar radii, they are termed as ICMEs or interplanetary coronal mass ejections [11]. The ICMEs are capable of interacting with earth's magnetosphere and can produce geomagnetic storms and cause power

surges in the grids which can potentially cause blackouts. A fast CME can reach the earth within hours while a slow CME could take days to do so.

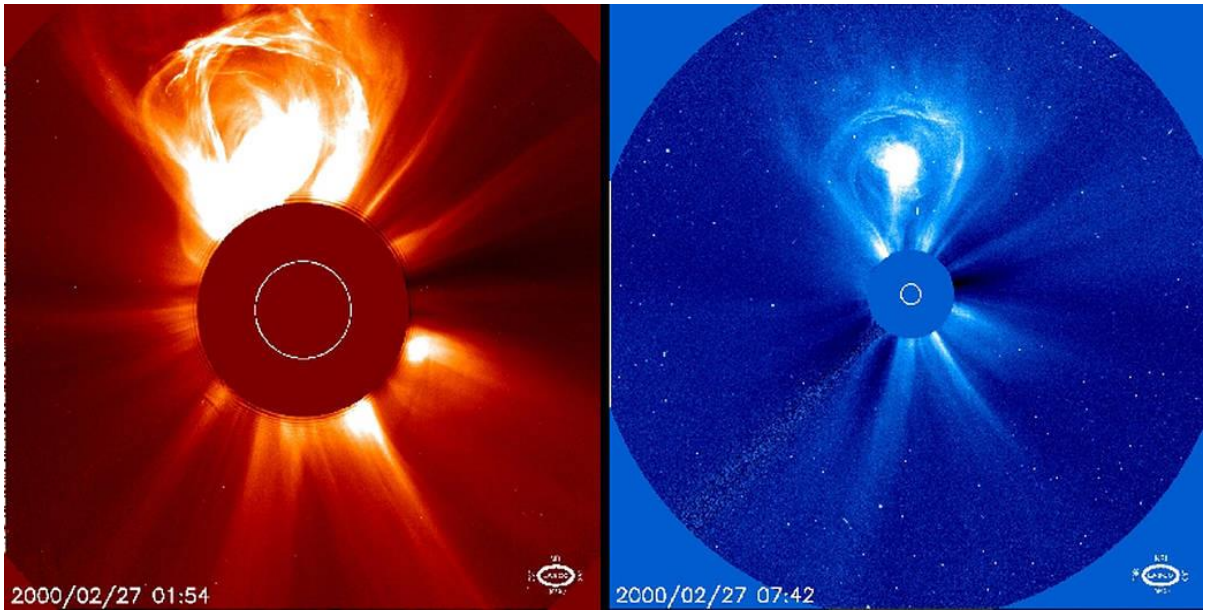


Figure 1.5: A coronal mass ejection on Feb. 27, 2000 taken by SOHO LASCO C2 and C3. Credit: SOHO ESA and NASA

Fig. 1.5 show a CME captured by Large Angle Spectroscopic Coronagraph (LASCO). Here we can clearly see the difference in field of view of both coronagraphs C2 and C3. The C2 has a field of view (fov) of 1.5 to 6 solar radii and C3 has a fov of 3.7 to 30 solar radii. A typical CME usually has a three part structure, a cavity of low electron density, a dense core, and a bright leading edge as shown in Fig. 1.6. The composition of CME remains uncertain. The answer lies in the complete understanding of the CME eruption mechanism. It makes sense that the material dominant in the launch zone would make up the majority of the CME's mass if the CME's launch close to the Sun is magnetically dominated.

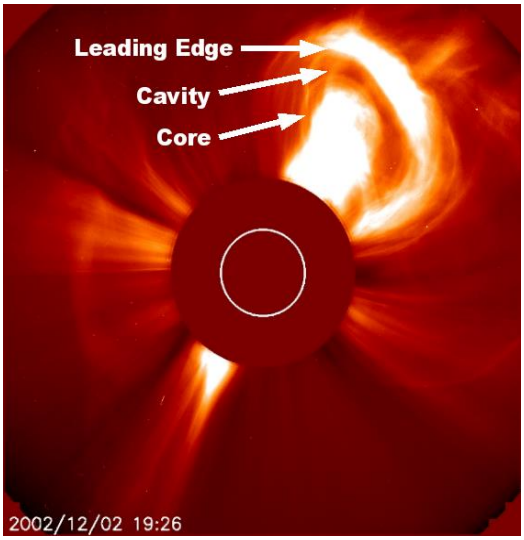


Figure 1.6: Image of a "classic" 3 part CME.

Image Courtesy: nasa.gov

1.5 Detection of CMEs

The basic predicament with observing CME and other coronal features is that the solar corona is faint as compared to the background solar disk. The bright background makes it extremely difficult to gaze at the features of the corona. The only time that the corona is normally visible is during a solar eclipse that completely obscures the sun's disc.

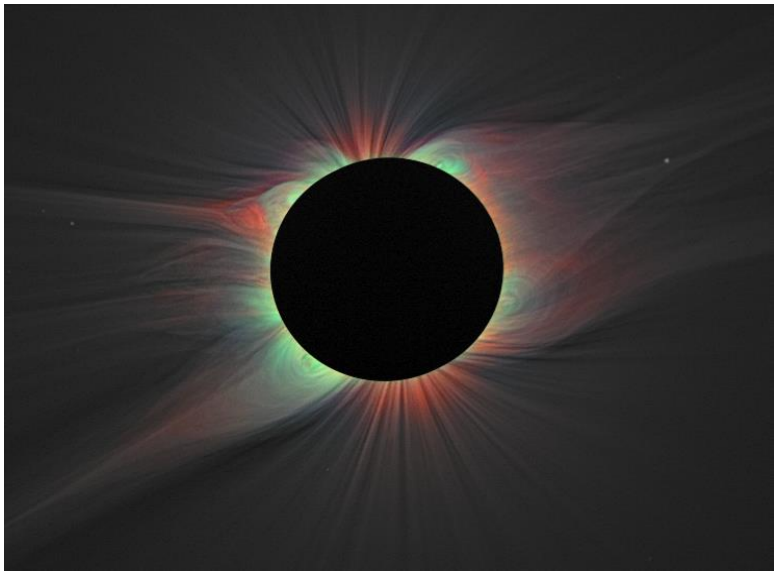


Figure 1.7: This image of the solar corona contains a color overlay of the emission from highly ionized iron lines and white light taken of the 2008 eclipse. Red indicates Fe XI 789.2 nm, blue represents Fe XIII 1074.7 nm, and green shows Fe XIV 530.3 nm.

Credit: *Habbal, et al*

Since solar eclipses are rare phenomenon and cannot always be dependable to observe the corona. Also the eclipse achieves totality for few minutes which is not ideal from an observational standpoint. Now, space-based coronagraphs are the primary method for detecting CMEs. They produce an artificial eclipse by obstructing most of the Sun's disk, revealing only the faint corona that surrounds it. The LASCO on board Solar and Heliospheric Observatory (SOHO), which has identified well over 10^4 CMEs since its debut in 1995. It has been the most effective coronagraph for CME detection to date. By studying the white light reflected by the electrons within the CME's plasma, LASCO can identify the CME (Thompson Scattering).

More subsequently, the LASCO coronagraph has been joined by other spacecraft-based coronagraphs. These are the COR coronagraphs, which function similarly to LASCO and are part of the Solar Terrestrial Relations Observatory (STEREO) satellite. Because of the two-dimensional nature of coronagraph images, CME images are actually projections into the sky plane. As a result, the appearance of a CME is greatly influenced by the direction of propagation. For instance, the identical CME with a component along the Sun-observer line might appear narrower and faster than one travelling very close to the sky's plane.

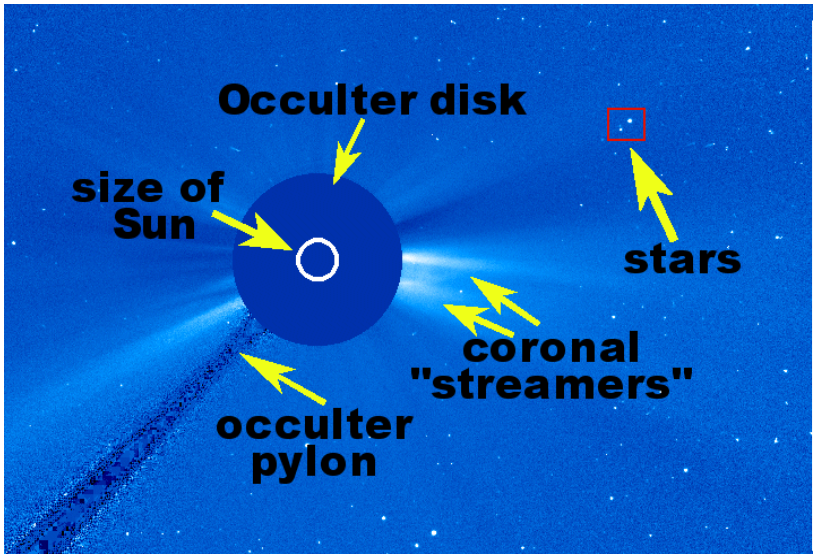


Figure 1.8: The graphic highlights the main characteristics of the LASCO C3image. The LASCO "C3" camera is the source of this specific picture. The "pylon," an arm that keeps the occulter in position, the solid occulter disc that simulates an artificial eclipse, and a white disc that represents the Sun was drawn on the occulter during image processing, are all visible.

Credits: lasco-www.nrl.navy.mil

The CMEs that are along the sun-observer line are called Halo CMEs. They are called halo because of the way they appear in the coronagraphic images. They appear to completely encircle the sun and have angular spread of 360º. Since Halo CMEs are directed towards the earth they are highly dangerous and much more geoeffective.

1.6 Need for detection of CMEs

The sun releases an incessant flow of plasma and magnetic field from its surface which slams the bodies in our solar system with particles and radiation called solar wind. It can propagate all the way to planetary surfaces unless thwarted by an atmosphere. It was first proposed in the 1950s by University of Chicago physicist Eugene Parker. Our earth is also enclosed within the solar wind and for the most part earth's magnetosphere shields us from it and the solar wind in turn protects us from the cosmic radiations.

Expulsions of CME cause ejection of plasma and magnetic field in the solar wind. Since our magnetosphere is sensitive to the composition and variations in the solar wind, this can cause geomagnetic disturbances. A magnetic reconnection is a temporary link between the magnetic fields of the Earth and an ICME. It can occur when an ICME enters the magnetosphere with a strong southward-directed magnetic field. This is seen in Fig. 1.9. The magnetic field lines open upon reconnection. By injecting the plasma from the CME directly into the weakened geomagnetic field, it exposes the Earth to the highly energetic particles. This can then cause geomagnetic storms, communication breakdowns, power outages, and serious damage to both space-based and ground-based equipment.

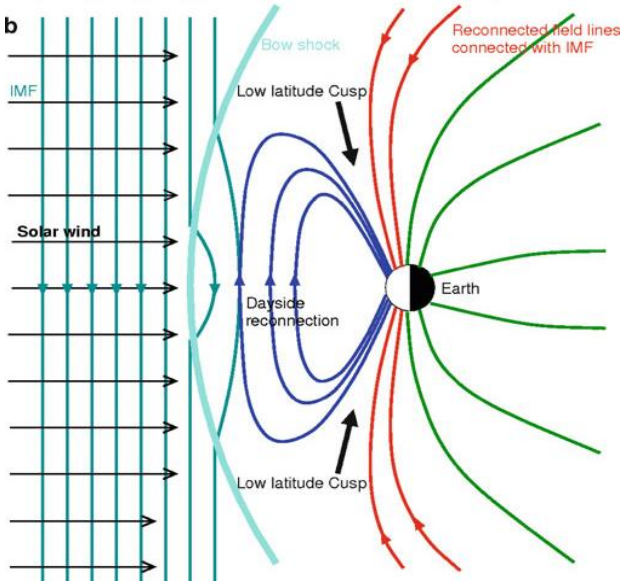


Figure 1.9: When a southward-directed magnetic field arrives at the Sun, magnetic reconnection enables more dayside field lines to open, moving the cusp region towards the equator resulting in an aurora being observed at lower latitudes

Effects of geomagnetic storms

1. Satellite electronics can be harmed by static-electric charge accumulation and discharge during magnetic storms.
2. High-altitude pilots and astronauts may be exposed to higher radiation doses.
3. Voltage spikes in the power grid can result in blackouts.
4. During storms, the ionosphere is heated and deformed, global positioning system (GPS) and communications might suffer, and long-range radio communication that depends on sub-ionospheric reflection may be difficult or impossible.

This pretext provides a clear and concise image that CMEs are critical drivers of the space weather. An earth directed CME can have devastating effects while the early detection of it can give us time to prepare which could be of outmost importance.

2. Hough Transforms

The Hough Transform is a feature isolation technique which is used to extract particular shapes within an image. It is used in image analysis, computer vision, and digital image processing. The feature to be tracked is to be defined by some parameters and then the image is transformed from the image space to the Hough or the parameter space. For example a straight line can have its slope and intercepts as its parameters. The object candidates are obtained as local maxima in the accumulator space following a voting method in a parameter space. Originally intended to identify geometric objects, the Hough transform is now also capable of identifying any arbitrary shape. Though the terms voting and accumulator may not be clear now but are explained in the later sections.

2.1 Linear Hough Transform

Linear Hough transform is the technique used to detect straight line, edges or ridges. So how do we do it? Any straight line in the Cartesian coordinate system/image plane is defined by equation y=m*x + c where m=slope and c =y intercept of the line as shown in figure below

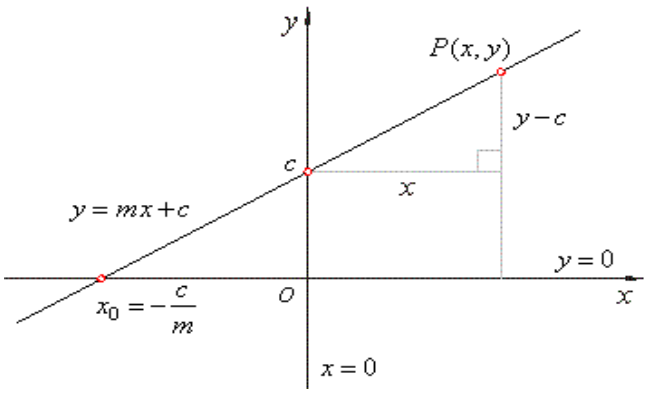


Figure 2.1: shows a straight line in the image space with the equation y=mx+c

Now we need to define some parameters which will be unique for any particular line. So the parameters of choice are r and theta (θ).

1. r which is the perpendicular distance of the line from the origin or a fixed point.
2. θ which is the angle made by the perpendicular to the line with respect to x axis calculated in anti-clock direction.

Now each line in image space is uniquely determined by a single point in the parameter space. Conversely each point in image space will correspond to infinite number of points in the Hough space as infinite number of lines can pass through each point and each line corresponds to a point. Since the transformation relation between the image space and the parameter space is given by $r=x\cos(\theta)+y\sin(\theta)$, the above mentioned infinite sequence of points will be of sinusoidal form. The infinite number of points can be avoided by quantizing your r and theta values on proper intervals.

So the question is how do we detect a straight line? It's quite simple. Suppose you have two or multiple points and you want to find a line passing through them. Remember that an infinite number of lines can pass through a point and since we are trying to find a line passing through all the points, that particular line should be common for all points. This is the secret of the detection algorithm. First we plot those points from image space to the Hough space which will give us a number of sinusoidal curves. Now the point where all these curves intersect will give us the value of r and theta of the line which passes through all the points and this will be our required line.

2.2 Algorithm

1. Initialize the values of ρ and θ. Usually the values range from [0,180] degrees for θ and [-d, d] for ρ where d is the length of the diagonal of the image. Alternatively you can also use [0,360] degrees for θ and [0, d] for ρ. Quantize the parameters accordingly.

2. Initialize a 2D accumulator array representing the Hough Space with dimension (n1, n2), where n1 is the number of ρ values and n2 is number of θ values and initialize all its values to zero. So each cell of the accumulator now points to a specific ρ and θ value.
3. For every pixel verify if it is an edge pixel. If yes, then, loop through all possible values of θ and calculate the corresponding ρ. Using the θ and ρ as index in the accumulator and increment the value at that position. So each time by incrementing we are basically establishing the fact that this particular line has been observed n number of times, where n is the value of a particular cell in accumulator array.
4. Go through all the values in the accumulator, if the local maxima in the accumulator are greater than a certain threshold, then that point represents a line in the image. Acquire the index of the maxima which is basically value of ρ and θ. Hence we have successfully detected the line with the parameters ρ and θ.

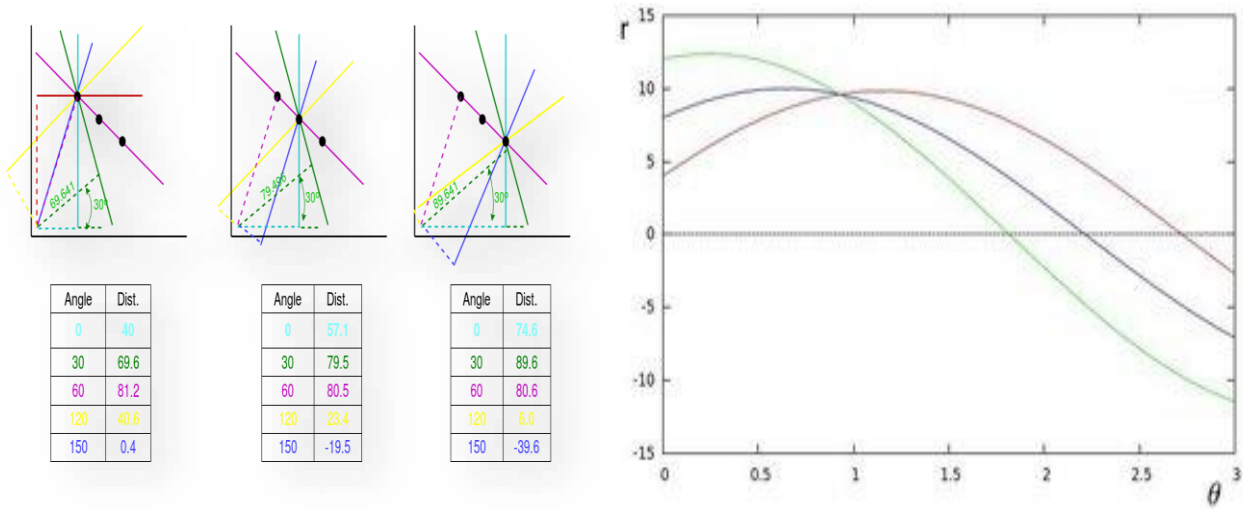


Figure 2.2: Hough Transform of 3 points from image space to Hough space.

In figure 2.2, figure on the left shows three points in the image space and number of lines that pass through each of them along with r and theta value of each line. The figure on the right shows those same three points in the parameter space and the point of intersection of the three curves corresponds to the line passing through all the three points with theta=60^{o} and r=80.76 units.

2.3 CACTUS

CACTUS is a software package for 'Computer Aided CME Tracking'. It automatically detects CMEs from the image sequences obtained from the LASCO onboard SOHO. The software provides a list of events, similar to classic catalogues such as CDAW [4], with angular width, principle angle, velocity, kinetic energy and mass estimations of the CMEs. In contrast to CDAW which is manually operated i.e. each event has to be checked by a human operator, these detections are faster and reduces human labour involved. Automated detection also reduces the subjectivity of the events as the detection criterion is written explicitly in a program.

URL for the CACTUS catalogue- https://www.sidc.be/cactus/

URL for the CDAW catalogue- http://cdaw.gsfc.nasa.gov/CME_list.

The aim of this work is to improve upon the previous detection algorithms being used in the automated catalogues. Here we try to use a different feature detection algorithm called the Generalized Hough Transform (GHT). Using this new approach we will try to upgrade the efficiency of the detection algorithms.

3. Generalized Hough Transform (GHT)

The generalized Hough transform, introduced by Dana H. Ballard in 1981 is a modification on the original Hough transforms as it has the ability to detect arbitrary shapes or the shapes with no geometric locus [5]. It needs a detailed description of the target object's precise shape. It uses the principle of template matching. Here we create a model (Hough model) of the arbitrary object we are trying to detect and then search for that model in our given image using the voting procedure similar to Hough transforms.

3.1 Procedure for GHT

Lets us go through the modus operandi of GHT one step at a time.

Step 1. Creating a model for the object to be detected.

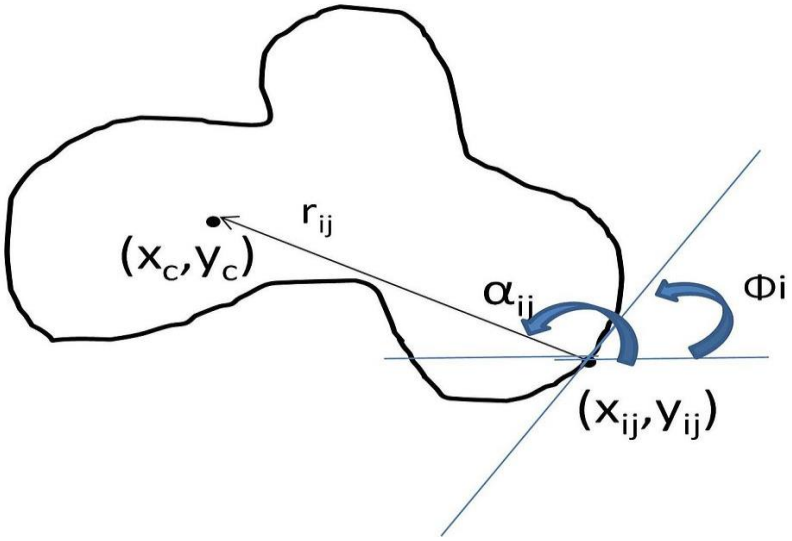


Figure 3.1: An arbitrary object for which we are trying to create a Hough model.

Suppose we have an arbitrary shaped object as shown in Fig. 3.1 for which we are trying to create a model.

1. Pick an anchor or reference point (x_c, y_c). Usually the average of the edge points is taken as the anchor points. Now we try to model the entire object around the anchor point. In layman's terms we are trying to see how the object looks from the perspective of the anchor point.
2. For a boundary point we calculate three things. Distance of the boundary point from the anchor point (r), orientation of vector r with respect to x axis (α), orientation of the boundary point (Φ). Now we store the tuple (r, α) as a function of Φ in a table called the Φ-table. We repeat this process for each boundary point.

   From the Fig X we can also derive a relation between (x_c, y_c) and (x, y) which turns out to

   $$x_c = x + r\cos(\alpha) \text{.....(3.1)}$$

   $$y_c = y + r\sin(\alpha) \text{.....(3.2)}$$
3. The Φ-table generated gives us the exact specification of our object with respect to the anchor points. The Φ-table would somewhat look as shown in Fig. 3.2

i	Φ	R=(r, α)
1	Φ_1	$^1R_1, ^1R_2, ^1R_3 \ldots\ldots\ldots\ldots ^1R_N$
2	Φ_2	$^2R_1, ^2R_2, ^2R_3 \ldots\ldots\ldots\ldots ^1R_M$
:	:	:
n	Φ_n	$^NR_1, ^NR_2, ^NR_3 \ldots\ldots\ldots\ldots ^NR_Z$

Figure 3.2: The phi table for an object will have the following appearance.

Step 2: Searching the model in the given image.

Our model is ready and now we try to match it in our given image, so we follow the given steps.

1. We loop through each pixel in the image and check if it is an edge point. For a pixel to be an edge point value of any of its RGB channel should be greater than 0 (for grayscale image all three channels will have the same value). For each detected edge point we calculate the orientation of the edge (Φ).
2. We then access the values of the R_n corresponding to that Φ value in the previously generated Φ table. Using the obtained values of r and α we calculate the anchor points using the equation (1) and (2) and increment their respective positions in the accumulator array. (So if the value in the x,y position of the accumulator array is 100, it testifies that the anchor points x,y have been voted or generated 100 times.)
3. The maxima in the accumulator array points to the presence of the anchor points which corresponds to the detection of an instance of the model.

I'd like to elaborate the last point. As we have previously seen that the Hough model is such that it creates a relation between the anchor point and the boundary points. This relation is such that each boundary point points to the anchor point. Now suppose in our image we have an object which is either exactly similar or similar to some degree with our Hough model, due to this similarity some or all points of the object will also point to a single point when the Hough routine is performed and global maxima will be achieved (assuming there is only one instance of the Hough model in the image). This maximum will solidify the fact that instance of the object in the image is present.

This is the detection procedure of the GHT

3.2 Algorithm

1. Quantize the parameter space: $P[X_{min} ... X_{max}] [Y_{min} \dots Y_{max}]$
2. Calculate the angle Φ for each edge point (x, y) and then salvage all the (r, α) values from the Φ -table.
3. For every (r, α) corresponding to the Φ value in the Φ-table, estimate the anchor points:

$$x_c = x + r\cos(\alpha), \quad y_c = y + r\sin(\alpha)$$

4. Increment the accumulator position with respect to the Anchor points obtained.($++P[x_c][y_c]$)
5. Local maxima in accumulator array P[x][y] denotes the possible detection of the object.
6. Index of the maxima is obtained and it denotes the location of the detected object.

One important step is the detection of the edge orientation. The orientation of a edge is detected by a inbuilt function called sobel operator. It uses two 3x3 kernels to approximate the gradients in x and y directions called G_x, G_y respectively. The direction of the gradient Θ at pixel (x, y) is:

$$\Theta = \text{atan}(G_y / G_x) \dots\dots\dots\dots (3)$$

The above mentioned algorithm has a flaw. It will not be able to detect the model in the image if it is either rotated by some degrees or uniformly scaled by some arbitrary factor. So for a more general case we introduce rotation and scaling feature to the preexisting algorithm.

We already know that

$$x_c = x + x' \qquad \text{where} \quad x' = r\cos(\alpha)$$

$$y_c = y + y' \qquad \text{where} \quad y' = r\sin(\alpha)$$

Now if the object has undergone rotation by angle θ and uniform scaling by a factor of s then

$$x_c = x - (x'\cos(\theta) - y'\sin(\theta))s \ldots\ldots..(3.4)$$

$$y_c = y - (x'\sin(\theta) + y'\cos(\theta))s \ldots\ldots..(3.5)$$

Now the model also accounts for each combination of the rotation angle θ and the scaling factor s, similar to before maxima will once again declare the detection of instance of the object. Addition of features makes the computation burden larger and the process becomes slow, so unless you have too many features or quantization of the parameters is very small, the program should run just fine.

An important tip- while defining your accumulator array, beware that its size should be greater than your image size. So often the value of the anchor points calculated by the algorithm may be larger than the image size, since it is just basic addition and an error will be forced. I have seen that taking 50 to 100 columns and rows greater than the size of the image does the job. If it does not then adjust accordingly.

3.3 Code for GHT

```
import cv2 as cv
import numpy as np
import matplotlib.pyplot as plt
im = cv.imread(r"path of the reference object")
im=im[:,:,0]
sobelxref = cv.Sobel(im,cv.CV_64F,1,0)
sobelyref = cv.Sobel(im,cv.CV_64F,0,1)
gradient_directionref = np.rad2deg(np.arctan2(sobelyref,sobelxref))
```

```
gradient_directionref = np.round(gradient_directionref,3)
table=np.zeros(((im.shape[0]*im.shape[1]),3))
image =cv.imread(r"path of the image in which object is to be detected")
image=image[:,:,0]
sobelx = cv.Sobel(image,cv.CV_64F,1,0)
sobely = cv.Sobel(image,cv.CV_64F,0,1)
gradient_direction = np.rad2deg(np.arctan2(sobely,sobelx))
gradient_direction = np.round(gradient_direction,3)
#calculating average of the edge points which will be used as anchor points
xsum=0
ysum=0
n=0
for x in range(im.shape[0]):
    for y in range(im.shape[1]):
        if im[x][y]!= 0:
            xsum+=x
            ysum+=y
            n+=1
xsum=int(xsum/n)
ysum=int(ysum/n)
print(xsum,ysum,n)
def findDistance(x1,y1):
    r = ((xsum-x1),(ysum-y1))
```

```
    return  r
n=0
for x in range(im.shape[0]):
    for y in range(im.shape[1]):
        if im[x][y]!= 0: # boundary point
            phi = gradient_directionref[x][y]
            r = findDistance(x,y)
            table[n][0]=phi
            table[n][1]=r[0]
            table[n][2]=r[1]
            n+=1
table=table[0:n,:]
print(table)
print('number of pixels in ref image',n)
plt.figure(figsize=(20,20))
plt.subplot(131),plt.imshow(im,"gray")
plt.subplot(132),plt.imshow(image,'gray')
plt.subplot(133),plt.imshow(acc,'gray')
print(np.amax(acc))
```

4. Results

As discussed in the previous section how the GHT works, here I will discuss the results of the algorithm. So basically models for the object were created and they were searched for in the image using the GHT and accumulator array was searched for maxima.

Trial 1

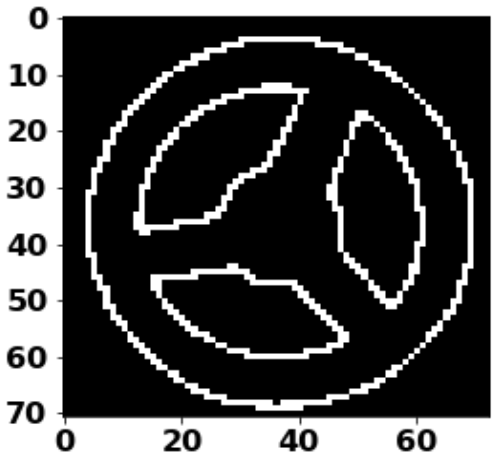


Figure 4.1: The image shows the object for which the Hough model is created in Trial 1.This image is referred as Model 1 in the text.

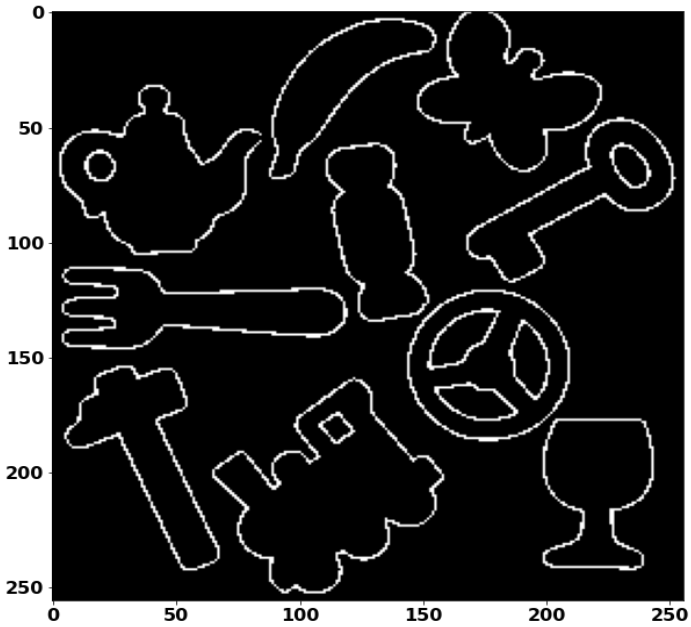


Figure 4.2: The figure shows the image in which the GHT is performed to detect the model. This is referred to as reference image in the text.

Model 1 shown in Fig. 4.1 has 544 edge points and after performing GHT on the reference image, a maxima of value 533 is observed in the accumulator array as shown in Fig. 4.3., which shows the detection of the object with accurate location (slightly right to the center of the image).

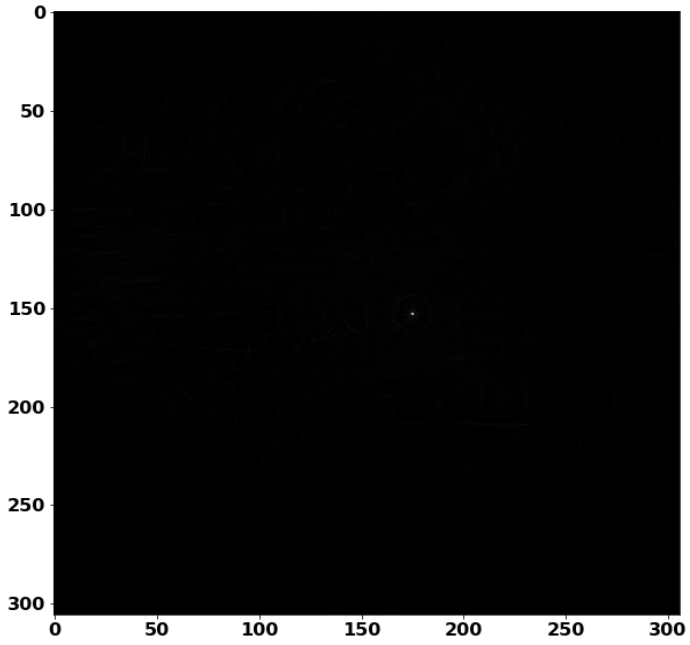


Figure 4.3: The image represents the accumulator array after searching for model 1 in the reference image. Object is detected at (150,175).

Trial 2

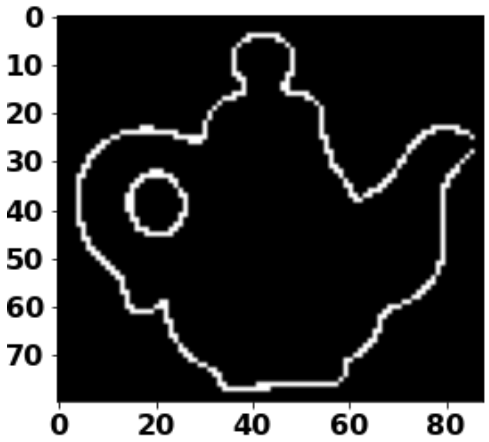


Figure 4.4: The image shows the object for which the Hough model is created in Trial 2.This image is referred as model 2 in the text.

In the second trial we use a new image to create a new Hough model called model 2, which is shown in Fig 4.4. Model 2 has 380 edge points and after performing GHT on the reference image, a maxima of value 380 is observed in the accumulator array as shown in Fig 4.5. It shows detection of model 2 with accurate location (top left side of the image)

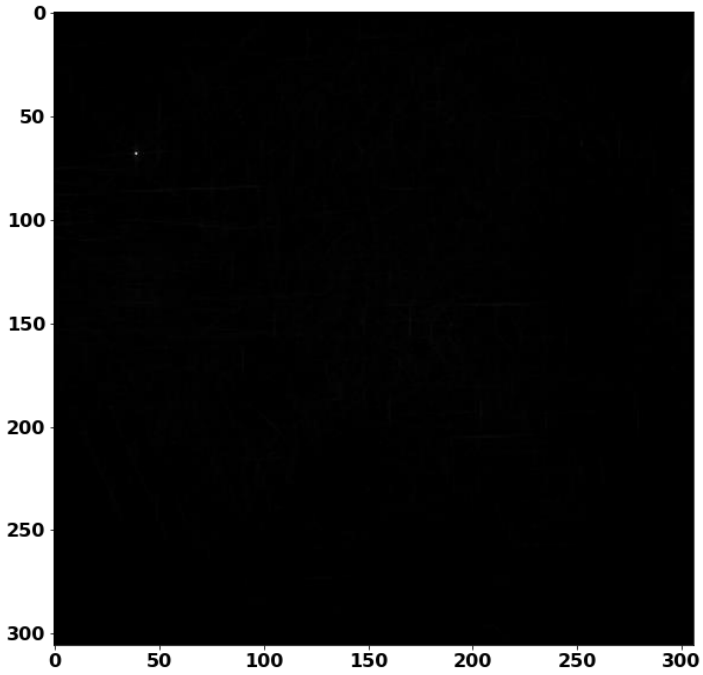


Figure 4.5: The image represents the accumulator array after searching for model 1 in the reference image. The object is detected at (65, 40).

5. Conclusion

The purpose of this work was to develop an automation technique to detect CMEs as the manual detection is subjective and requires lot of manpower. Chapter 1 discusses the basics of Sun slowly leading to the informatics on CMEs. Then the next chapters discuss Hough Transform, a feature detection technique. The subsequent chapter discusses the Generalized Hough Transform technique which we successfully used to detect objects in chapter four.

As evident in the fourth chapter that GHT is a pretty neat technique to detect arbitrary objects, provided that we have a model ready for them. The problem that is likely to occur with CMEs is that they don't have a fixed shape or structure. This problem could be solved in one of two ways. First we introduce new features to the Hough transform which will grant us flexibility in terms of the shape of the object but as the number of features increases so does the computational load. So we can only add features up to a certain limit without making the algorithm sluggish. Secondly we can make large number of models of different kinds of CMEs and subsequently a adaptive thresholding technique can be used to detect the CMEs adequately.

As far as detection technique is concerned, it is fairly evident that GHT is a efficient technique and still has vast aptitude for application.

References

1. Duda, R.O.; Hart, P. E. (January 1972). "Use of the Hough Transformation to Detect Lines and Curves in Pictures". *Comm. ACM*. **15**: 11–15. doi:10.1145/361237.361242
2. Generalized Hough transform, Ballard and Brown, section 4.3.4, Sonka et al., section 5.2.6
3. Gopalswamy, N., Yashiro, S., Michalek, G. et al. The SOHO/LASCO CME Catalog. Earth Moon Planet **104,** 295–313 (2009). https://doi.org/10.1007/s11038-008-9282-7
4. Howard T. (2011), Coronal Mass Ejections: An Introduction, Astrophysics and Space Science Library, VOL. 376, doi:10.1007/978-1-4419-8789-1
5. Solar Astrophysics by Peter V. Foukal, ISBN 3-527-40374-4
6. Sun Fact Sheet, NASA Goddard Flight Center, https://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html
7. Swinburne University of Technology https://astronomy.swin.edu.au/cosmos/m/main+sequence+lifetime
8. The Heliopedia, Nasa https://www.nasa.gov/mission_pages/sunearth/the-heliopedia
9. The Sun An introduction, Dr. Michael Stix, ISBN 978-3-642-62477-3 ISBN 978-3-642-56042-2 (eBook) DOI 10.1007/978-3-642-56042-2
10. V. Pant, S. Willems, L. Rodriguez, M. Mierla, D. Banerjee, and J. A. Davies (Dec 2016) Automated detection of coronal mass ejections in stereo heliospheric imager data doi:10.3847/1538-4357/833/1/80
11. Webb D. and Howard T. (2012), Coronal Mass Ejections: Observations, Living Rev.

Solar Phys., VOL. 9, 3

12. Yashiro, S.; Gopalswamy, N.; Michalek, G.; Cyr, O. C. St.; Plunkett, S. P.; Rish, N. B.; Howard, R. A. (July 2004). "A catalog of white light coronal mass ejections observed by the SOHO spacecraft". Journal of Geophysical Research: Atmospheres. **109** (A7). Bibcode:2004JGRA..109.7105Y. doi:10.1029/2003JA010282